Backward error of polynomial eigenvalue problems solved by linearization of Lagrange interpolants

Piers W. Lawrence
Robert M. Corless

Report TW 655, February 2015

KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Backward error of polynomial eigenvalue problems solved by linearization of Lagrange interpolants

Piers W. Lawrence
Robert M. Corless

Report TW 655, February 2015

Department of Computer Science, KU Leuven

Abstract

This article considers the backward error of the solution of polynomial eigenvalue problems expressed as Lagrange interpolants. One of the most common strategies to solve polynomial eigenvalue problems is to linearize, which is to say that the polynomial eigenvalue problem is transformed into an equivalent larger linear eigenvalue problem, and solved using any appropriate eigensolver. Much of the existing literature on the backward error of polynomial eigenvalue problems focuses on polynomials expressed in the classical monomial basis. Hence, the objective of this article is to carry out the necessary extensions for polynomials expressed in the Lagrange basis. We construct one-sided factorizations that give simple expressions relating the eigenvectors of the linearization to the eigenvectors of the polynomial eigenvalue problem. Using these relations, we are able to bound the backward error of an approximate eigenpair of the polynomial eigenvalue problem relative to the backward error of an approximate eigenpair of the linearization. We develop bounds for the backward error involving both the norms of the polynomial coefficients and the properties of the Lagrange basis generated by the interpolation nodes. We also present numerous numerical examples to illustrate the numerical properties of the linearization, and develop a balancing strategy to improve the accuracy of the computed solutions.

Keywords : Stability, backward error, polynomial eigenvalue problem, linearization, Lagrange interpolation, barycentric formula
MSC : Primary : 65F15, 15A18

BACKWARD ERROR OF POLYNOMIAL EIGENVALUE PROBLEMS SOLVED BY LINEARIZATION OF LAGRANGE INTERPOLANTS*

PIERS W. LAWRENCE† AND ROBERT M. CORLESS‡

Abstract. This article considers the backward error of the solution of polynomial eigenvalue problems expressed as Lagrange interpolants. One of the most common strategies to solve polynomial eigenvalue problems is to linearize, which is to say that the polynomial eigenvalue problem is transformed into an equivalent larger linear eigenvalue problem, and solved using any appropriate eigensolver. Much of the existing literature on the backward error of polynomial eigenvalue problems focuses on polynomials expressed in the classical monomial basis. Hence, the objective of this article is to carry out the necessary extensions for polynomials expressed in the Lagrange basis. We construct one-sided factorizations that give simple expressions relating the eigenvectors of the linearization to the eigenvectors of the polynomial eigenvalue problem. Using these relations, we are able to bound the backward error of an approximate eigenpair of the polynomial eigenvalue problem relative to the backward error of an approximate eigenpair of the linearization. We develop bounds for the backward error involving both the norms of the polynomial coefficients and the properties of the Lagrange basis generated by the interpolation nodes. We also present numerous numerical examples to illustrate the numerical properties of the linearization, and develop a balancing strategy to improve the accuracy of the computed solutions.

Key words. Stability; backward error; polynomial eigenvalue problem; linearization; Lagrange interpolation; barycentric formula

AMS subject classifications. 65F15; 15A18

1. Introduction.

In this article, we are interested in the solution of polynomial eigenvalue problems (PEPs) expressed in barycentric Lagrange form. In these PEPs, polynomials are defined by their values $P_j = P(\sigma_j)$ at a set of $n+1$ distinct nodes $\{\sigma_0, \ldots, \sigma_n\}$. Throughout this article, we will use the first form of the barycentric interpolation formula [6] (also known as the modified Lagrange formula [26]), defined by

$$P(\lambda) = \ell(\lambda) \sum_{j=0}^{n} \frac{\beta_j P_j}{\lambda - \sigma_j} , \tag{1.1}$$

where the degree of $P(\lambda)$ is at most n, the Lagrange basis coefficients are $P_i \in \mathbb{C}^{m\times m}$, and the node polynomial $\ell(\lambda)$ and barycentric weights β_j are defined by

$$\ell(\lambda) = \prod_{i=0}^{n} (\lambda - \sigma_i) , \quad \beta_j = \prod_{\substack{k=0 \\ k\neq j}}^{n} (\sigma_j - \sigma_k)^{-1} , \quad 0 \leq j \leq n . \tag{1.2}$$

The usual Lagrange basis polynomials $\ell_j(\lambda)$ can be written in terms of the node polynomial and barycentric weights as

$$\ell_j(\lambda) = \ell(\lambda) \frac{\beta_j}{(\lambda - \sigma_j)} , \quad 0 \leq j \leq n . \tag{1.3}$$

*This research was partially supported by the Research Council KU Leuven, project OT/10/038 ("Multi-parameter model order reduction and its applications"), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). The scientific responsibility rests with its authors.

†Department of Computer Science, KU Leuven, 3001 Leuven (Heverlee), Belgium (piers.lawrence@cs.kuleuven.be)

‡Department of Applied Mathematics, The University of Western Ontario, London, ON N6A 5B7, Canada (rcorless@uwo.ca)

The polynomial eigenvalue problem is to find scalars λ and nonzero vectors x and y that satisfy $P(\lambda)x = 0$ and $y^*P(\lambda) = 0$. In this work, we shall also assume that the polynomial is regular, that is, $\det P(\lambda)$ is not identically zero.

In the literature, polynomial eigenvalue problems are most commonly expressed in the monomial basis [23, 28, 40]. However, there has been growing interest in PEPs expressed in other bases [2, 15, 41], either due to the construction of the polynomials themselves, or in order to take advantage of the properties of a particular polynomial basis.

2. Linearization. One of the most widespread solution methods for solving PEPs is to linearize [22], which is to say they are transformed into larger generalized eigenvalue problems having the same eigenstructure. That is, the linearization has the same eigenvalues as the polynomial, and the eigenvectors of the polynomial are easily recovered from the eigenvectors of the linearization. Since the problem is now a linear generalized eigenvalue problem, one may use any of the well-established algorithms for computing the eigenvalues and eigenvectors of the linearization, for example, the QZ algorithm [36].

Certainly, linearization has proven to be an extremely convenient method to compute all of the roots of scalar polynomials, and many different linearizations have been proposed. Almost all of the linearizations proposed in the literature to date are constructed using the monomial basis coefficients of the polynomial [4, 13, 19, 34], although there have been some notable exceptions for polynomials satisfying three term recurrence relations [5, 24]. Most of the aforementioned linearizations were developed for computing the roots of scalar polynomials only. However, almost all can be extended to matrix polynomials in a very simple way, for example, the extensions proposed by Amiraslani et al. [2], as well as the generalization of the Fiedler companion forms proposed by Antoniou and Vologiannidis [4].

Before introducing the particular linearizations for PEPs expressed in barycentric Lagrange form, we will first introduce the basic definitions of linearization and strong linearization relevant to the discussion. This is a restatement of the definition of linearization introduced by Gohberg, Kaashoek, and Lancaster [22], and later named strong linearization by Lancaster and Psarrakos [29].

DEFINITION 2.1. *Linearization of order mn [22]: A linear matrix pencil $\mathcal{L}(\lambda) = \lambda B - A$ is said to be a linearization of $P(\lambda)$ of order mn if $\mathcal{L}(\lambda)$ is of size $mn \times mn$ and the polynomials $P(\lambda)$ and $\mathcal{L}(\lambda)$ are related in the following way:*

$$E(\lambda)\mathcal{L}(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{m(n-1)} \end{bmatrix}, \tag{2.1}$$

where $E(\lambda)$ and $F(\lambda)$ are some $mn \times mn$ invertible matrix polynomials with non-vanishing determinant independent of λ.

DEFINITION 2.2. *Strong linearization (of order mn) [22, 28, 29]: A matrix pencil $\mathcal{L}(\lambda) = \lambda B - A$ is said to be a strong linearization (of order mn) if it is a linearization of order mn and there exist two unimodular polynomial matrices $M(\lambda)$ and $N(\lambda)$ such that*

$$M(\lambda)\mathcal{L}^{\#}(\lambda)N(\lambda) = \begin{bmatrix} P^{\#}(\lambda) & 0 \\ 0 & I_{m(n-1)} \end{bmatrix}, \tag{2.2}$$

where $\mathcal{L}^{\#}(\lambda) = B - \lambda A$, and where $P^{\#}(\lambda) = \lambda^n P(1/\lambda)$ is known as the reverse polynomial [29], or the n-reversal [14, Def. 2.12]. Note that the choice of n does not need to coincide with the degree, that is, we admit $n \geq \deg P(\lambda)$.

REMARK 1. *According to Lancaster and Psarrakos [29], the term strong linearization was so named by Gohberg, Kaashoek, and Lancaster [22]. In that work, they do not make the restriction that* $n = \deg P(\lambda)$*, that is, they allow for some of the leading coefficients of the polynomial to be equal to zero. This fact was also pointed out by Lancaster and Psarrakos in the note [29], and evidently needed further clarification [28] by making the explicit distinction that the polynomials in question have extended degree, that is, when* $n > \deg P(\lambda)$ *(see, for example, De Terán and Dopico [12]). Yet another approach was taken by De Terán, Dopico, and Mackey [14] where they distinguish such polynomials by employing the grade of the polynomial.*

In this work, we examine polynomials that are either already expressed in the Lagrange basis or have been transformed into the Lagrange basis by sampling the polynomial where it is well conditioned to do so. For the Lagrange basis, it appears that the first linearization (of order $m(n+2)$) to be described for matrix polynomials was the arrowhead linearization proposed by Corless [9] (it was recently made aware to the authors that in an earlier work of Fiedler [18], a similar arrowhead construction for finding the zeros of scalar polynomials was proposed). We work with a slightly different form of the linearization, defined by $\mathcal{L}(\lambda) = \lambda B - A$, where

$$A = \begin{bmatrix} 0 & -P_0 & \cdots & -P_n \\ \beta_0 I & \sigma_0 I & & \\ \vdots & & \ddots & \\ \beta_n I & & & \sigma_n I \end{bmatrix}, \quad B = \begin{bmatrix} 0 & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix}. \tag{2.3}$$

Using Schur's determinant formula, we easily see that $\det \mathcal{L}(\lambda) = \det P(\lambda)$, and thus the eigenvalues of the linearization coincide with those of the polynomial $P(\lambda)$. Furthermore, Amiraslani et al. [2] demonstrated that $\mathcal{L}(\lambda)$ is a strong linearization of order $m(n+2)$. That is, the linearization $\mathcal{L}(\lambda)$ has the same finite and infinite eigenvalues of $P(\lambda)$ as well as an additional $2m$ spurious infinite eigenvalues.

Since the linearization (2.3) has an additional $2m$ eigenvalues at infinity, there have been efforts to construct linearizations of Lagrange interpolants with smaller dimension [42]. We show in §4 that these smaller linearizations can be formed by simple equivalence transformations applied to (2.3); thus, the methodology developed in this manuscript for deriving bounds for the ratios of the backward errors is also applicable to these linearizations. Although the bounds would be different from the ones we develop here in this work.

3. Backward errors. We will first introduce some general definitions and notation for backward errors of the solutions of polynomial eigenvalue problems, and then develop bounds for the backward error of approximate eigenpairs of $P(\lambda)$ relative to the backward error of the approximate eigenpairs of the linearization $\mathcal{L}(\lambda)$. These bounds provide useful information as to when the eigenvalues of $P(\lambda)$ can be computed with small backward errors.

3.1. Definitions and notation. Throughout this article, we are primarily interested in the normwise backward error of an approximate eigenpair of the polynomial $P(\lambda)$. We would like to know the extent to which we need to perturb the original polynomial coefficients in order for an approximate eigenpair to be the exact solution of the perturbed problem.

The normwise backward error of a finite approximate right eigenpair (λ, x) of a polynomial $P(\lambda)$ is defined by

$$\eta_P(\lambda, x) = \min\{\varepsilon : (P(\lambda) + \Delta P(\lambda))x = 0,\ \|\Delta P_j\|_2 \le \varepsilon \|P_j\|_2,\ 0 \le j \le n\}, \tag{3.1}$$

where $\Delta P(\lambda) = \sum_{j=0}^{n} \Delta P_j \ell_j(\lambda)$, and the $\ell_j(\lambda)$'s are the Lagrange basis polynomials (1.3). Similarly, for an approximate left eigenpair (λ, y^*) of $P(\lambda)$, the normwise backward error is defined by

$$\eta_P(\lambda, y^*) = \min\{\varepsilon : y^*(P(\lambda) + \Delta P(\lambda)) = 0,\ \|\Delta P_j\|_2 \leq \varepsilon \|P_j\|_2,\ 0 \leq j \leq n\}\,. \quad (3.2)$$

These definitions are straightforward generalizations of the definitions for the monomial basis [27]. However, the difference here is that we consider perturbations to the Lagrange basis coefficients P_j rather than to the monomial coefficients.

For the monomial basis, Tisseur [38] obtained explicit expressions for the backward errors of the approximate left and right eigenpairs of $P(\lambda)$, (λ, y^*) and (λ, x), respectively, given by

$$\eta_P(\lambda, x) = \frac{\|P(\lambda)x\|_2}{B_M(\lambda)\|x\|_2}\,, \quad \eta_P(\lambda, y^*) = \frac{\|y^*P(\lambda)\|_2}{B_M(\lambda)\|y\|_2}\,. \quad (3.3)$$

In these expressions, $B_M(\lambda)$ is defined by

$$B_M(\lambda) = \sum_{j=0}^{n} \|A_j\|_2 |\lambda|^j\,, \quad (3.4)$$

where the A_j's are the monomial basis coefficients.

Amiraslani [1] and Corless et al. [11] have also extended this result to the Lagrange basis by considering the ε-pseudospectrum of polynomials expressed in other bases. The equivalent expressions (for the 2-norm) for the backward errors are obtained by replacing $B_M(\lambda)$ in (3.3) by $B_L(\lambda)$, where

$$B_L(\lambda) = \sum_{j=0}^{n} \|P_j\|_2 |\ell_j(\lambda)|\,. \quad (3.5)$$

It has also been noted by both Tisseur and Higham [39] and Corless et al. [11] that the norms $\|A_j\|_2$ and $\|P_j\|_2$ occurring in (3.4) and (3.5), respectively, can be replaced by nonnegative weights α_j, not all equal to zero, which control how the perturbations to the coefficients are measured. Farouki and Rajan [17] have also defined these quantities as (absolute) condition numbers for the evaluation of polynomials (see also Corless and Fillion [10, Thm. 2.8, pp. 63]). For this work, we only consider relative perturbations in the coefficient matrices, and thus we use the expressions (3.4) and (3.5) above.

By applying the expressions in (3.3) to the linear pencil $\mathcal{L}(\lambda)$, the backward errors of approximate left and right eigenpairs, (λ, u^*) and (λ, v), are given by

$$\eta_{\mathcal{L}}(\lambda, u^*) = \frac{\|u^*\mathcal{L}(\lambda)\|_2}{(|\lambda|\|B\|_2 + \|A\|_2)\,\|u\|_2}\,, \quad \eta_{\mathcal{L}}(\lambda, v) = \frac{\|\mathcal{L}(\lambda)v\|_2}{(|\lambda|\|B\|_2 + \|A\|_2)\,\|v\|_2}\,, \quad (3.6)$$

where $\mathcal{L}(\lambda) = \lambda B - A$, and the coefficient matrices A and B are defined in (2.3).

3.2. Backward error of $P(\lambda)$ relative to $\mathcal{L}(\lambda)$. The main objective of this article is to bound the backward errors of eigenpairs of $P(\lambda)$ relative to the backward errors of eigenpairs of $\mathcal{L}(\lambda)$ for polynomials expressed in the Lagrange basis. To achieve this goal, we need to find relations between the eigenvectors of the polynomial $P(\lambda)$ and those of the linearization $\mathcal{L}(\lambda)$, where the polynomial and linearization are

defined in (1.1) and (2.3), respectively. We may utilize the framework developed by Grammont et al. [25] to analyze nonlinear eigenvalue problems; once we find appropriate one-sided factorizations of the linearization $\mathcal{L}(\lambda)$, we are immediately able to obtain relations for the backward error of the polynomial eigenvalue problem relative to the backward error of the linearization.

Thus, our goal is to determine the polynomial matrices $G(\lambda)$ and $H(\lambda)$, which satisfy

$$G(\lambda)\mathcal{L}(\lambda) = g^T \otimes P(\lambda)\,, \quad \mathcal{L}(\lambda)H(\lambda) = h \otimes P(\lambda)\,, \tag{3.7}$$

and where $G(\lambda)$ and $H(\lambda)^T$ have dimension m by $m(n+2)$, $g \in \mathbb{C}^{n+2}$, and $h \in \mathbb{C}^{n+2}$. By direct computation, we easily see that

$$H(\lambda) = \Lambda(\lambda) \otimes I = \begin{bmatrix} \ell(\lambda) \\ \ell_0(\lambda) \\ \vdots \\ \ell_n(\lambda) \end{bmatrix} \otimes I \tag{3.8}$$

satisfies the second relation in (3.7) with $h = e_1$, the first unit vector, since

$$\mathcal{L}(\lambda)H(\lambda) = \begin{bmatrix} 0 & P_0 & \cdots & P_n \\ -\beta_0 I & (\lambda-\sigma_0)I & & \\ \vdots & & \ddots & \\ -\beta_n I & & & (\lambda-\sigma_n)I \end{bmatrix} \begin{bmatrix} \ell(\lambda)I \\ \ell_0(\lambda)I \\ \vdots \\ \ell_n(\lambda)I \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{n} \ell_i(\lambda)P_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.9}$$

It is only slightly more complicated to construct a suitable polynomial matrix $G(\lambda)$, and one may easily verify that

$$G(\lambda) = \begin{bmatrix} \ell(\lambda)I & -\ell_0(\lambda)P_0\beta_0^{-1} & \cdots & -\ell_n(\lambda)P_n\beta_n^{-1} \end{bmatrix} \tag{3.10}$$

satisfies the first relation in (3.7) with $g = e_1$. Now that we have found two specific one-sided factorizations of the linearization, we use them to relate the eigenvectors of $\mathcal{L}(\lambda)$ to those of $P(\lambda)$ in the following lemma.

LEMMA 3.1. *Suppose that (λ, u^*) and (λ, v) are left and right approximate eigenpairs of $\mathcal{L}(\lambda)$, respectively. Then, $y^* = u^*(e_1 \otimes I)$ and $x = (e_1^T \otimes I)v$ are left and right eigenvectors of $P(\lambda)$, respectively, provided that they are both nonzero.*

Proof. From the relations (3.7), it follows that

$$u^*\mathcal{L}(\lambda)H(\lambda) = u^*(e_1 \otimes P(\lambda)) = u^*(e_1 \otimes I)P(\lambda)\,, \tag{3.11}$$

and

$$G(\lambda)\mathcal{L}(\lambda)v = (e_1^T \otimes P(\lambda))v = P(\lambda)(e_1^T \otimes I)v\,. \tag{3.12}$$

Provided that $y^* = u^*(e_1 \otimes I)$ and $x = (e_1^T \otimes I)v$ are nonzero, these equations imply that $y^*P(\lambda) = 0$ and $P(\lambda)x = 0$ by virtue of u and v being eigenvectors of $\mathcal{L}(\lambda)$, and thus y and x are eigenvectors of $P(\lambda)$ corresponding to eigenvalue λ. ∎

REMARK 2. *We need to consider the possibility of encountering cases where $y^* = u^*(e_1 \otimes I)$ or $x = (e_1^T \otimes I)v$ could be equal to zero. It is a happy coincidence that*

this occurs precisely when an eigenvalue coincides with one of the interpolation nodes σ_i. Hence, we consider $\mathcal{L}(\sigma_i)$, which has a zero block on the diagonal where $(\sigma_i - \sigma_i)I$ and in this block column we have P_i. Since $P_i = P(\sigma_i)$ is singular, an approximate eigenvector of the linearization is zero everywhere except for the corresponding block, that is, from the $(i+2)$nd block.

Now, having found suitable one-sided factorizations, and the appropriate relations between eigenvectors of the linearization and the polynomial, we state our first theorem relating the backward error of the approximate eigenpairs of $P(\lambda)$ to those of $\mathcal{L}(\lambda)$.

THEOREM 3.2. *The backward error $\eta_P(\lambda, x)$ of an approximate right eigenpair, (λ, x) of $P(\lambda)$, relative to the backward error $\eta_{\mathcal{L}}(\lambda, v)$ of a corresponding approximate right eigenpair, (λ, v) of $\mathcal{L}(\lambda)$, can be bounded by*

$$\frac{\eta_P(\lambda, x)}{\eta_{\mathcal{L}}(\lambda, v)} \leq \frac{|\lambda|\|B\|_2 + \|A\|_2}{B_L(\lambda)} \cdot \frac{\|G(\lambda)\|_2 \|v\|_2}{\|x\|_2}\,, \tag{3.13}$$

where $G(\lambda)$ is defined in (3.10)*, and $\mathcal{L}(\lambda) = \lambda B - A$. We may similarly bound the backward error of an approximate left eigenpair, (λ, y^*) of $P(\lambda)$, by*

$$\frac{\eta_P(\lambda, y^*)}{\eta_{\mathcal{L}}(\lambda, u^*)} \leq \frac{|\lambda|\|B\|_2 + \|A\|_2}{B_L(\lambda)} \cdot \frac{\|H(\lambda)\|_2 \|u\|_2}{\|y\|_2}\,, \tag{3.14}$$

where $\eta_{\mathcal{L}}(\lambda, u^*)$ is the backward error of a corresponding approximate left eigenpair (λ, u^*) of $\mathcal{L}(\lambda)$, and $H(\lambda)$ is defined in (3.8)*.*

Proof. Our proof follows the discussion in Higham et al. [27, §2.2]. The only difference is that we use $B_L(\lambda)$, since the polynomials are expressed in the Lagrange basis. In essence, we begin with the definitions of the backward errors of eigenpairs of $P(\lambda)$ (3.3), then through the relations (3.11) and (3.12) we can write them in terms of the linearization $\mathcal{L}(\lambda)$, the one sided factorizations, and the eigenvectors of the linearization. After which, we combine these expressions with (3.6) to arrive at the desired upper bound. ∎

Now that we have established a relationship between the backward error of $P(\lambda)$ and the backward error of $\mathcal{L}(\lambda)$, we would like to know what conditions need to be satisfied in order to obtain small backward errors in the solution of the PEP. Thus, we investigate the conditions under which the ratios $\eta_P(\lambda, x)/\eta_{\mathcal{L}}(\lambda, v)$ and $\eta_P(\lambda, y^*)/\eta_{\mathcal{L}}(\lambda, u^*)$ are approximately equal to one. Small backward errors in the computation of the approximate eigenvalues of $\mathcal{L}(\lambda)$ will then necessarily lead to small backward errors in the solution of the PEP.

THEOREM 3.3. *Let (λ, v) be an approximate eigenpair of $\mathcal{L}(\lambda)$. If x is recovered from the first m rows of v, we obtain the bound*

$$\frac{\eta_P(\lambda, x)}{\eta_{\mathcal{L}}(\lambda, v)} \leq \frac{(|\lambda| + \|A\|_2)\max\left(1, \max_i \|P_i\|_2 |\beta_i^{-1}|\right)\|\Lambda(\lambda)\|_2}{\sum_{i=0}^{n} \|P_i\|_2 |\ell_i(\lambda)|} \cdot \frac{\|v\|_2}{\|x\|_2}\,, \tag{3.15}$$

where $\Lambda(\lambda) = \begin{bmatrix} \ell(\lambda) & \ell_0(\lambda) & \cdots & \ell_n(\lambda) \end{bmatrix}^T$, the node polynomial $\ell(\lambda)$ is defined in (1.2)*, and the $\ell_j(\lambda)$'s are the Lagrange basis polynomials. Similarly, let (λ, u^*) be an approximate left eigenpair of $\mathcal{L}(\lambda)$. If y^* is recovered from the first m columns of u^*, we obtain the bound*

$$\frac{\eta_P(\lambda, y^*)}{\eta_{\mathcal{L}}(\lambda, u^*)} \leq \frac{(|\lambda| + \|A\|_2)\,\|\Lambda(\lambda)\|_2}{\sum_{i=0}^{n} \|P_i\|_2 |\ell_i(\lambda)|} \frac{\|u\|_2}{\|y\|_2}\,. \tag{3.16}$$

Proof. For an approximate left eigenpair (λ, u^*), we combine (3.14) with (3.8). Given that $\|H(\lambda)\|_2 = \|\Lambda(\lambda) \otimes I\|_2 = \|\Lambda(\lambda)\|_2$, we immediately obtain the upper bound (3.16). For an approximate right eigenpair (λ, v), we combine (3.13) with (3.10), and rewrite $G(\lambda) = \Lambda(\lambda)\,\mathrm{diag}\left\{I, -P_0\beta_0^{-1}, \ldots, -P_n\beta_n^{-1}\right\}$, where $\mathrm{diag}\,\{\cdot\}$ constructs a block diagonal matrix. Thus, $\|G(\lambda)\|_2 \leq \|\Lambda(\lambda)\|_2 \max\left(1, \max_i \|P_i\|_2|\beta_i^{-1}|\right)$, and we obtain the upper bound (3.15). ∎

From the bounds (3.16) and (3.15), we begin to see the conditions under which the backward error of the solution of the PEP is not that much larger than the backward error of the linearization.

REMARK 3. *The influence of the choice of nodes manifests itself in the term* $\|\Lambda(\lambda)\|_2$. *This term behaves essentially like the Lebesgue function, and hence we would ideally like to choose sets of nodes that give small Lebesgue functions. However, we also need points that are good approximations of the eigenvalues themselves, since we also need to limit the effect of* $\ell(\lambda)$, *the first entry of* $\Lambda(\lambda)$. *We will not be able to satisfy this condition for all mn eigenvalues, but if we are interested in obtaining some eigenvalues with small backward errors, it is clear that placing a well-conditioned set of nodes close to the eigenvalues of interest will achieve this goal. Moreover, the norms* $\|P_i\|_2$ *should not be too large, and when the magnitudes of all of the polynomial coefficients are approximately equal to one, we minimize the upper bound.*

3.3. Balanced linearizations. In this section, we discuss the issue of balancing the linearization $\mathcal{L}(\lambda)$ so as to improve the backward errors of the computed eigenpairs of $P(\lambda)$. Lawrence has shown [32] that balancing linearizations of scalar polynomials can lead to significant gains in the accuracy of computed roots. Indeed, for standard eigenvalue problems, balancing a matrix prior to computing its eigenvalues has become a standard technique [3]. However, the landscape is not so well established for generalized and polynomial eigenvalue problems. For generalized eigenvalue problems, Lemonnier and Van Dooren [33] suggest a balancing strategy to bring the pencil closer to some standardized normal pencil. For the polynomial eigenvalue problems, the most commonly used technique appears to be eigenvalue parameter scaling, exemplified by the works of Fan, Lin, and Van Dooren [16] for the quadratic eigenvalue problem, and by Gaubert and Sharify [21] for the higher degree polynomials based on tropical roots. Further, Higham, Li, and Tisseur [27] propose a diagonal scaling that modifies only the identity blocks in the companion matrix in order to improve the backward errors of computed eigenpairs.

In this work, we apply block diagonal similarity transformations to the linearization in order to improve the ratios $\eta_P(\lambda, x)/\eta_{\widehat{\mathcal{L}}}(\lambda, \widehat{v})$ and $\eta_P(\lambda, y^*)/\eta_{\widehat{\mathcal{L}}}(\lambda, \widehat{u}^*)$, where $\widehat{\mathcal{L}}(\lambda) = D_s^{-1}\mathcal{L}(\lambda)D_s$ is the balanced linearization, D_s is a block diagonal matrix, and $\widehat{u}$ and $\widehat{v}$ are approximate left and right eigenvectors of $\widehat{\mathcal{L}}(\lambda)$, respectively. From the one sided factorizations (3.7) with $g = h = e_1$, and if the first diagonal block of D_s is equal to the identity we obtain

$$(G(\lambda)D_s)(D_s^{-1}\mathcal{L}(\lambda)D_s) = e_1^T \otimes P(\lambda), \quad (D_s^{-1}\mathcal{L}(\lambda)D_s)(D_s^{-1}H(\lambda)) = e_1 \otimes P(\lambda). \tag{3.17}$$

Thus, $\widehat{G}(\lambda) = G(\lambda)D_s$ and $\widehat{H}(\lambda) = D_s^{-1}H(\lambda)$ are one-sided factorizations of $\mathcal{L}(\lambda)$. Applying Theorem 3.2 to the balanced linearization $\widehat{\mathcal{L}}(\lambda)$, we may simply replace the matrices involved in the one sided factorization by their balanced versions, and replace the eigenvectors of $\mathcal{L}(\lambda)$ with those of $\widehat{\mathcal{L}}(\lambda)$. Furthermore, we can form the

upper bounds

$$\frac{\eta_P(\lambda,x)}{\eta_{\widehat{\mathcal{L}}}(\lambda,\widehat{v})} \leq \frac{|\lambda|+\|\widehat{A}\|_2}{B_L(\lambda)} \cdot \frac{\|\Lambda(\lambda)\|_2 \|D_G D_s\|_2 \|\widehat{v}\|_2}{\|x\|_2}\,, \tag{3.18}$$

and

$$\frac{\eta_P(\lambda,y^*)}{\eta_{\widehat{\mathcal{L}}}(\lambda,\widehat{u}^*)} \leq \frac{|\lambda|+\|\widehat{A}\|_2}{B_L(\lambda)} \cdot \frac{\|\Lambda(\lambda)\|_2 \|D_s^{-1}\|_2 \|\widehat{u}\|_2}{\|y\|_2}\,, \tag{3.19}$$

where $\widehat{A} = D_s^{-1} A D_s$, and $D_G = \operatorname{diag}\left\{I, -P_0\beta_0^{-1}, \ldots, -P_n\beta_n^{-1}\right\}$. Thus, we suggest the following heuristic diagonal scaling matrix

$$D_s = \operatorname{diag}\left\{I, I\sqrt{|\beta_0|/\|P_0\|_2}, \ldots, I\sqrt{|\beta_n|/\|P_n\|_2}\right\}. \tag{3.20}$$

The reasoning behind this is that we then have $\|D_s^{-1}\|_2 = \|D_G D_s\|_2$, equal to

$$\|D_s\|_2 = \max\left(1, \max_i \sqrt{\frac{\|P_i\|_2}{|\beta_0|}}\right), \tag{3.21}$$

and thus, both upper bounds are the same apart from the terms $\|\widehat{v}\|_2/\|x\|_2$ and $\|\widehat{u}\|_2/\|y\|_2$. Furthermore, if we examine the term $\|\widehat{A}\|_2$ closely, we see that the norms of the off diagonal blocks are made equal, that is, we see that $|\widehat{\beta}_i| = \|\widehat{P}_i\|_2$ for $0 \leq i \leq n$, where $\widehat{\beta}_i$ and $\widehat{P}_i$ are the off diagonal entries in the balanced matrix $\widehat{A}$. The consequence of this is that the matrix A is then balanced (blockwise) in the sense of Parlett and Reinsch [37], that is the above scaling solves the optimization problem

$$\inf_D \|D^{-1} A D\|_F\,, \tag{3.22}$$

where D is constrained to have block diagonal entries and its first block entry is equal to the identity.

REMARK 4. *It could so happen that $\|P_i\|_2 = 0$ for some i. However, this would mean that σ_i is an eigenvalue of multiplicity m, and we should perform a similarity to move this block of eigenvalues to the top of the linearization where they can be decoupled from the problem.*

4. Related linearizations and removal of infinite eigenvalues. Recently, Van Beeumen et al. [42] have developed new linearizations for the Lagrange basis that have smaller dimension than (2.3), in order to achieve one-to-one correspondence between the eigenvalues of the PEP and the eigenvalues of the linearization. In this section, we show how the spurious infinite eigenvalues of the linearization (2.3) can be decoupled via constant equivalence transformations, and thus we recover the same linearizations as those proposed by Van Beeuman et al. [42].

In order to decouple the spurious infinite eigenvalues from the linearization (2.3), we bring the linearization to generalized Hessenberg form. One way to do this is to apply a sequence of block Gauss transformations on the right of $\mathcal{L}(\lambda)$, annihilating the β_j's in the first block column. We first define the ratios $\theta_i = \beta_{i-1}/\beta_i$, and explicitly form the equivalence transformation matrix

$$E_\theta = \begin{bmatrix} 0 & I & & & \\ I & 0 & & & \\ 0 & I & -\theta_1 I & & \\ & & \ddots & \ddots & \\ & & & I & -\theta_n I \end{bmatrix}. \tag{4.1}$$

Once this matrix is applied to $\mathcal{L}(\lambda)$, we obtain the following equivalent linearization:

$$E_\theta\mathcal{L}(\lambda) = \left[\begin{array}{c|cccc} -\beta_0 I & (\lambda-\sigma_0)I & 0 & & \\ \hline 0 & P_0 & P_1 & \cdots & P_n \\ 0 & (\lambda-\sigma_0)I & -(\lambda-\sigma_1)\theta_1 I & & \\ & & \ddots & \ddots & \\ & & & (\lambda-\sigma_{n-1})I & -(\lambda-\sigma_n)\theta_n I \end{array}\right]. \tag{4.2}$$

From this equation, we see immediately that we may deflate a block of m infinite eigenvalues. We thus arrive at the same $(n+1)m$ by $(n+1)m$ linearization proposed by Van Beeumen et al. [42, Thm. 4.4] in the lower right block.

Let us now point out a unitary alternative to the equivalence transformation E_θ to decouple the infinite eigenvalues. The method was developed for scalar polynomial linearizations by Lawrence [32], and is straightforwardly extended to the matrix case. The method is also related to a procedure for the construction of vector orthogonal polynomials [8]. The construction produces unitary matrix Q_1 such that

$$Q_1^* \begin{bmatrix} \beta_0 & \sigma_0 & \\ \vdots & & \ddots \\ \beta_n & & & \sigma_n \end{bmatrix} Q = \begin{bmatrix} \alpha e_1 & H \end{bmatrix}, \tag{4.3}$$

where H is an upper Hessenberg matrix, and $Q = \operatorname{diag}\{1, Q_1\}$. If we now form $\widetilde{\mathcal{L}}(\lambda) = (Q^* \otimes I_m)\mathcal{L}(\lambda)(Q \otimes I_m)$, we see that

$$\widetilde{\mathcal{L}}(\lambda) = \begin{bmatrix} 0 & \widetilde{P} \\ -\alpha I & \lambda I - (H \otimes I_m) \end{bmatrix}, \tag{4.4}$$

where $\widetilde{P} = P(Q_1 \otimes I_m)$, and where $P = \begin{bmatrix} P_0 & \cdots & P_n \end{bmatrix}$. Furthermore, by permuting first two block rows of $\widetilde{\mathcal{L}}(\lambda)$ yields the following equivalent linearization

$$\left[\begin{array}{c|ccc} -\alpha I & (\lambda-h_{0,0})I & -h_{0,1}I & \cdots & -h_{0,n}I \\ \hline 0 & \widetilde{P}_0 & \widetilde{P}_1 & \cdots & \widetilde{P}_n \\ 0 & -h_{1,0}I & (\lambda-h_{1,1})I & \vdots & -h_{0,n}I \\ & & \ddots & \ddots & \vdots \\ & & & -h_{n,n-1}I & (\lambda-h_{n,n})I \end{array}\right]. \tag{4.5}$$

One of the advantages to this approach comes when we aim to deflate more infinite eigenvalues of the pencil, since we can do this via unitary transformations on the linearization (4.5). Suppose that $\widetilde{P}_0 = U_0 S_0 V_0^*$ is the singular value decomposition of $\widetilde{P}_0$, then multiplying the lower right block of (4.5) by $U^* = \operatorname{diag}\{U_0^*, V_0^*, \ldots, V_0^*\}$ on the left, and $V = I_{n+1} \otimes V_0$ on the right. Now, if there are any more infinite eigenvalues in addition to the $2m$ ones introduced by the linearization, theses will be detected as the zero singular values of $\widetilde{P}_0$ and can be removed easily, as shown by Lawrence [31]. If there are no additional infinite eigenvalues, then we can use unitary rotations to annihilate the block $-h_{1,0}I$ and decouple the second set of m infinite eigenvalues.

5. Numerical examples. In this section, we illustrate the backward error of computing eigenpairs of $P(\lambda)$ via the linearization $\mathcal{L}(\lambda)$ and also via the balanced linearization $\widehat{\mathcal{L}}(\lambda)$. The examples are taken from a variety of sources, some of which are available in the collection NLEVP [7]. The polynomials are expressed in the monomial basis, and thus we first need to sample the polynomials at a set of $n+1$ interpolation nodes, where n is the degree of the polynomial. The linearization $\mathcal{L}(\lambda)$ is constructed from these samples and the computed barycentric weights. We also construct the balanced linearization $\widehat{\mathcal{L}}(\lambda)$ discussed in §3.3. The generalized eigenvalues and the left and right eigenvectors of the linearization are computed in MATLAB using the function `qz`. Throughout this section, we use (3.3) for computing the backward errors of eigenpairs of $P(\lambda)$, and (3.6) for computing the backward errors of eigenpairs $\mathcal{L}(\lambda)$ and $\widehat{\mathcal{L}}(\lambda)$ (replacing the coefficients with the balanced versions). For the computation of the upper bounds, we use the expressions (3.15) and (3.16) for the unbalanced linearization, and (3.18) and (3.19) for the balanced linearization.

5.1. Butterfly. Our first example is available in the NLEVP collection [7], proposed by Mehrmann and Watkins [35]. The polynomial is a 64 by 64 quartic with T-even structure. The spectrum has a butterfly shape. In the monomial basis, the polynomial is given by

$$P(\lambda) = \lambda^4 A_4 + \lambda^3 A_3 + \lambda^2 A_2 + \lambda A_1 + A_0 \,, \tag{5.1}$$

where A_4, A_2, and A_0 are real symmetric matrices. The matrices A_3 and A_1 are real skew-symmetric.

Because the polynomial is given in monomial form, we sample the polynomial at five Chebyshev points of the second kind, on the interval $[-1, 1]$. The computed eigenvalues are shown in Figure 5.1; they show good visual agreement to the eigenvalues computed in [35]. The distribution of the ratios of the backward errors

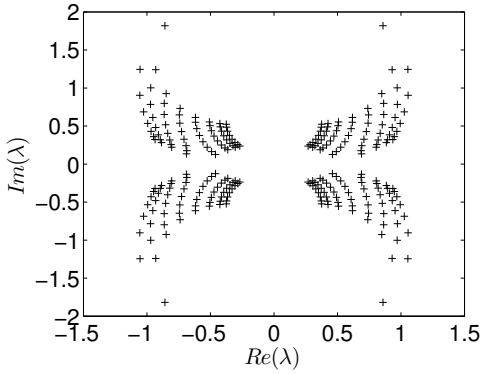
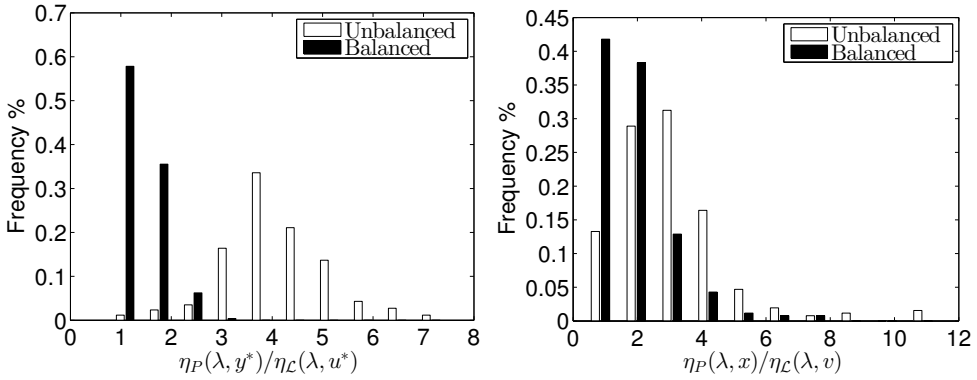


Fig. 5.1: Butterfly example—eigenvalue distribution

$\eta_P(\lambda, x)/\eta_{\mathcal{L}}(\lambda, v)$ and $\eta_P(\lambda, y^*)/\eta_{\mathcal{L}}(\lambda, u^*)$ are shown in Figure 5.2, along with the same ratios for the balanced linearization . The backward errors of the eigenpairs of the linearization are all of the order of the machine precision $\varepsilon_M \approx 1.1 \times 10^{-16}$, and thus we are ensured that the backward errors of the eigenpairs of the PEP are not much larger. We also see that balancing the linearization reduces the ratios of the backward errors significantly: almost half of the eigenvalues have ratios approximately equal to one for both the left and the right eigenpairs. Furthermore, as we



(a) Left eigenpair (b) Right eigenpair

Fig. 5.2: Butterfly example—backward error ratio distributions

show in Table 5.1, the upper bounds (3.16), (3.15), (3.18), and (3.19) approximate the ratios fairly well. We also see an improvement in the upper bounds for the balanced linearization In these computations, the maxima and minima are computed over all

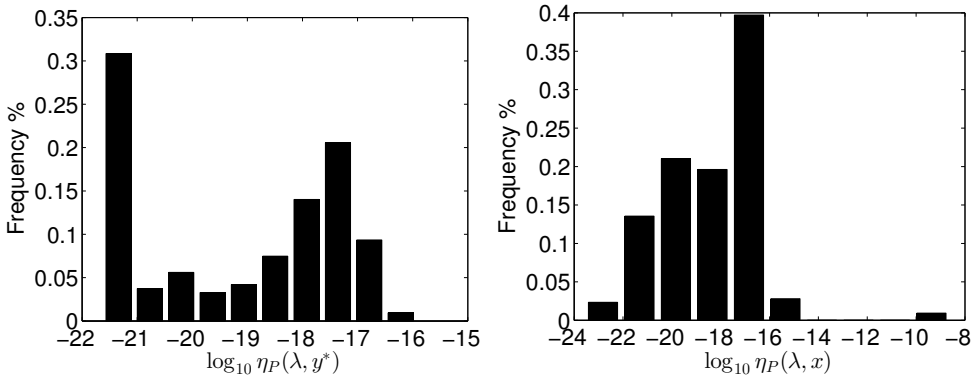
Table 5.1: Butterfly example—bound comparison

	Eigenpair	$\min \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound	$\max \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound
Unbalanced:	Left	1.22	5.47	7.49	18.3
	Right	0.406	27.6	11.4	239
Balanced:	Left	0.748	5.58	2.88	14.1
	Right	0.293	8.26	7.12	65.4

eigenvalues, and the upper bound corresponds to those eigenvalues.

5.2. Speaker enclosure. Our second example is also taken from the NLEVP collection [7]. The polynomial is the quadratic $P(\lambda) = \lambda^2 M + \lambda C + K$, where $M, C, K \in \mathbb{C}^{107\times 107}$, arising from a finite element model of a speaker enclosure. There is a large variation in the norms of the monomial basis coefficients: $\|M\|_2 = 1$, $\|C\|_2 = 5.7 \times 10^{-2}$, and $\|K\|_2 = 1 \times 10^7$.

We interpolate $P(\lambda)$ at the nodes $\{-i, 0, i\}$. At these nodes, $\|P_j\|_2 \approx 1 \times 10^7$, and so we have already, in a sense, equalized the norms of the coefficients through interpolation. The linearization $P(\lambda)$ is then balanced using the strategy described in §3.3, and this further equalizes the norms of the blocks P_j and the barycentric weights β_j. All the computed eigenvalues of the balanced linearization have real parts equal to zero, with the exception of the double eigenvalue at zero. This was not expected, since the linearization involves complex non-symmetric matrices. The logarithms of the backward errors of the eigenvalues of $P(\lambda)$ are shown in Figure 5.3. We show these values rather than the ratio of the backward errors because the backward errors of the eigenvalues of the linearization are all $O(10^{-15})$ in magnitude. The backward errors of the eigenpairs of $P(\lambda)$ are excellent, with the exception of the two eigenvalues close to zero for the right eigenvectors. However, the error bounds (3.16) and (3.15) do not predict the small backward error of the eigenpairs of $P(\lambda)$ relative to those of $\mathcal{L}(\lambda)$.



(a) Left eigenpair (b) Right eigenpair

Fig. 5.3: Speaker enclosure example—backward error distributions

We show the maximum ratios obtained in Table 5.2, together with the corresponding upper bounds. It would appear that the QZ algorithm is able to take advantage of the

Table 5.2: Speaker enclosure—upper bound comparison

	Eigenpair	$\min \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound	$\max \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound
Unbalanced:	Left	3.6×10^{-7}	1.04×10^{4}	6.55	5.58×10^{3}
	Right	3.59×10^{-6}	2.52×10^{11}	1.7×10^{10}	6.46×10^{11}
Balanced:	Left	1.29×10^{-6}	2.92×10^{4}	0.37	2.00
	Right	3.3×10^{-10}	7.2×10^{4}	1.3×10^{6}	4.9×10^{7}

structure of the linearization in some way, since eigenvalues with zero real parts are produced. However, we do not have a concrete explanation as to why this behaviour occurs.

5.3. Damped mass-spring system. The third example we investigate is a connected damped mass-spring system described by both Higham et al. [27] and Tisseur and Meerbergen [40, §3.9]. The polynomial $P(\lambda)$ is a 100 by 100 quadratic $P(\lambda) = \lambda^2 M + \lambda C + K$, where: $M = I$; C is tridiagonal, with super- and subdiagonal elements all equal to -64 and diagonal elements equal to $128, 192, 192, \ldots, 192, 128$; and K is tridiagonal, with super- and subdiagonal elements all equal to -1 and diagonal elements equal to $2, 3, 3, \ldots, 3, 2$. All of the eigenvalues are real and negative, 50 of which range from -320 to -64, while the remaining 50 are all approximately equal to -1.56×10^{-2}. We interpolate $P(\lambda)$ at the nodes $\{-0.01, 0, 0.01\}$. The eigenvalues of $P(\lambda)$ are all real, and hence we plot the real part against the index of the eigenvalue, as shown in Figure 5.4. Figure 5.5 shows the distribution of the logarithm of the backward errors of the eigenvalues of $P(\lambda)$ computed using the scaled linearization $\widehat{\mathcal{L}}(\lambda)$. Both figures show a distinct separation between two groups of eigenvalues. The eigenvalues near -1.56×10^{-2} are computed with backward errors of $O(10^{-12})$, whereas the larger magnitude eigenvalues are all computed with backward errors of $O(10^{-16})$. This behaviour has also been observed by Higham et al. [27], where the

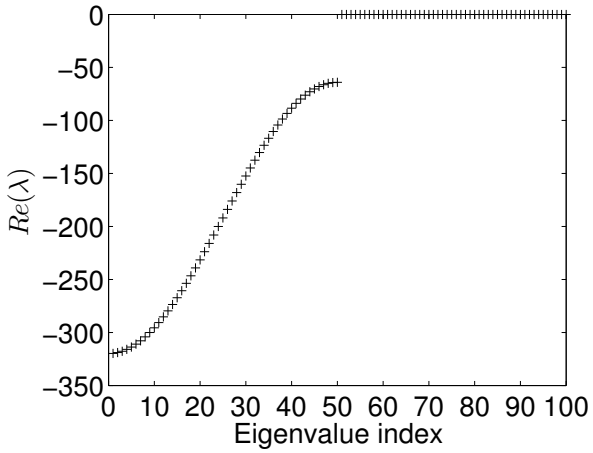
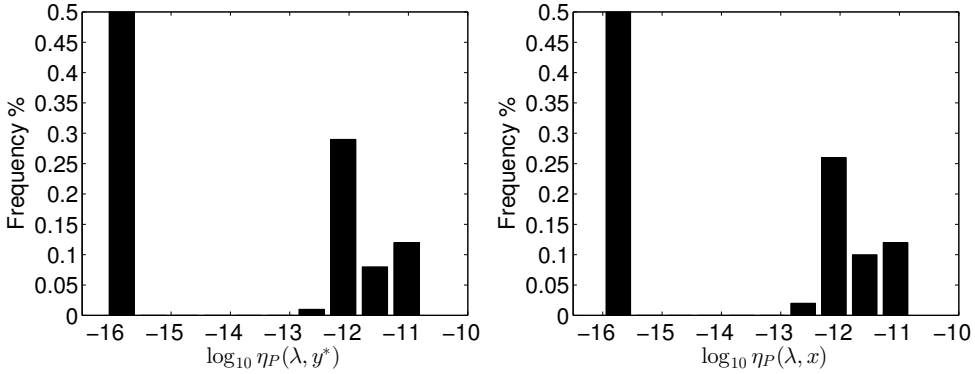


Fig. 5.4: Damped mass spring system—eigenvalue distribution



(a) Left eigenpair (b) Right eigenpair

Fig. 5.5: Damped mass spring system—backward error distributions

backward errors can only be small for one of the two groups of eigenvalues. We compare the upper bounds for the ratios of these backward errors in Table 5.3, where we additionally compare the maximum and minimum ratios and upper bounds for the set of eigenvalues of larger magnitude. Although we see a considerable reduction in the ratios and the upper bounds for the balanced linearization, the overestimation of the upper bound is still roughly two orders of magnitude. Part of the reasoning for the large backward errors of the small magnitude eigenvalues is the block from which the eigenvectors of $P(\lambda)$ are recovered from those of $\mathcal{L}(\lambda)$. For example, we see from (3.9) that we could recover the right eigenvector x from any of the blocks of v; this is a similar situation in the monomial basis [27], where the eigenvalues are recovered from the block of v having the largest norm.

5.4. Damped gyroscopic system. For our final example, we examine the damped gyroscopic system proposed in [30]. The polynomial $P(\lambda)$ is constructed

Table 5.3: Damped mass spring system—upper bound comparison

	Eigenpair	$\min \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound	$\max \frac{\eta_P}{\eta_{\mathcal{L}}}$	Upper bound
Unbalanced:	Left	31.5	1.54×10^3	3.81×10^5	1.09×10^6
	Right	6.68	2.53×10^5	5.13×10^5	2.37×10^9
Balanced:	Left	0.113	66.2	1.77×10^4	1.24×10^6
	Right	0.13	110	1.01×10^4	1.29×10^6
	Left ($\lambda \leq -63$)	0.113	66.2	0.293	110
	Right ($\lambda \leq -63$)	0.13	110	0.329	218

as follows: let N denote the 10 by 10 nilpotent matrix having ones on the subdiagonal and zeros elsewhere, and let I denote the 10 by 10 identity matrix. Define $\widehat{M} = (4I + N + N^T)/6$, $\widehat{G} = N - N^T$, and $\widehat{K} = N + N^T - 2I$. Then define the matrices M, G, and K, using the Kronecker product $\otimes$, by

$$\begin{aligned} M &= I \otimes \widehat{M} + 1.3\widehat{M} \otimes I \,, \\ G &= 1.35I \otimes \widehat{G} + 1.1\widehat{G} \otimes I \,, \\ K &= I \otimes \widehat{K} + 1.2\widehat{K} \otimes I \,. \end{aligned}$$

The damping matrix D is tridiagonal with super- and subdiagonal elements equal to -0.1 and diagonal elements equal to 0.2. The quadratic polynomial $P(\lambda)$ we examine is defined by

$$P(\lambda) = \lambda^2 M + \lambda(G + D) + K \,.$$

We interpolate $P(\lambda)$ at the nodes $\{-1.8, 0, 1.8\}$. In addition to the eigenvalues and backward errors, we also compute the weighted ε-pseudospectrum (see, for example, [30, 39]), shown together with the eigenvalues in Figure 5.6. The dotted line represents where the absolute condition numbers for evaluation are equal for the Lagrange basis and the monomial basis, that is $B_M(\lambda) = B_L(\lambda)$. Within the dotted line, the condition number of the Lagrange basis is somewhat smaller than that of the monomial basis, and hence we can expect to compute more accurate eigenvalues there.

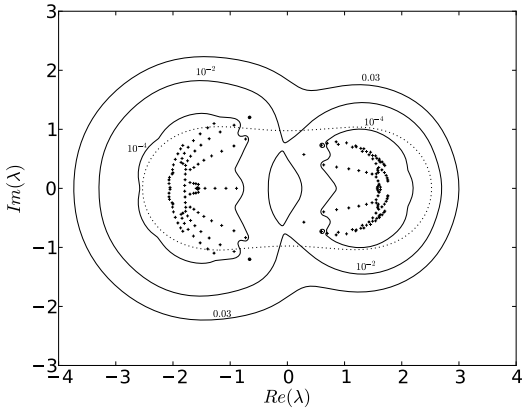
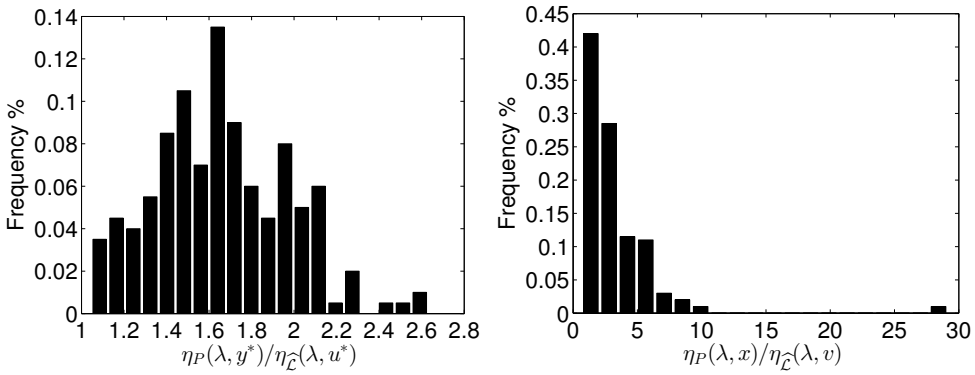


Fig. 5.6: Damped gyroscopic system—eigenvalue distribution and pseudospectra (the dotted line represents the level curve where $B_M(\lambda) = B_L(\lambda)$)

Furthermore, because we are able to choose the locations of the nodes, we can ensure that eigenvalues of interest are computed accurately by placing nodes near the eigenvalues. If nothing is known about the spectrum of $P(\lambda)$, then we may initially compute the eigenvalues using, for example, Chebyshev nodes on the interval $[-1, 1]$. We may then interpolate $P(\lambda)$ using some of the computed eigenvalues as nodes. This kind of iterative algorithm has been used successfully in the scalar case [20], and we expect to obtain similar results in the matrix case. For the monomial basis, we have no such flexibility.

The ratios of the backward errors are shown in Figure 5.7 for the balanced linearization $\widehat{\mathcal{L}}(\lambda)$. We see that for the left eigenpairs, the ratios are close to one. The ratios are not so favourable for the right eigenpairs, and there are two outliers close to 30. These two eigenvalues are the closest ones to the node at -1.8, and we suspect that the backward error could be improved by choosing a different block from which to recover the eigenvector of $P(\lambda)$.



(a) Left eigenpair (b) Right eigenpair

Fig. 5.7: Damped gyroscopic system—backward error distributions

We also compute the maximum ratios of the backward error and the corresponding upper bound. These values are shown in Table 5.4, where we see that for the left eigenvectors, the bounds are quite reasonable. Again, we also see the upper bound overestimate the ratios for the right eigenpairs by about an order of magnitude for the balanced linearization.

Table 5.4: Damped gyroscopic system—upper bound comparison

	Eigenpair	$\min \frac{\eta_P}{\eta_\mathcal{L}}$	Upper bound	$\max \frac{\eta_P}{\eta_\mathcal{L}}$	Upper bound
Unbalanced:	Left	4.7	17.5	18.8	26.5
	Right	3.91	272	58.7	321
Balanced:	Left	1.05	9.84	2.63	9.62
	Right	0.671	18.2	29.1	95.9

6. Concluding remarks. In this article, we have investigated the backward error of the solution to polynomial eigenvalue problems expressed in the Lagrange basis, solved via linearization. We have derived upper bounds for the ratio of the backward error of eigenpairs of the polynomial to those of the linearization. The conditions under which these ratios are close to one depend strongly upon the interpolation nodes used, as well as on the norms of the polynomial coefficients. In the Lagrange basis setting, the polynomial coefficients are the values of the polynomial at the nodes. Thus, in order to have good backward errors, we are guided to choose nodes that give polynomial coefficients with norm close to one, in conjunction with having a set of nodes that gives a well-conditioned basis. We have described a block-wise balancing strategy for the linearization, and this balancing can significantly improve the backward errors of the computed eigenpairs.

REFERENCES

[1] A. Amiraslani, *New Algorithms for Matrices, Polynomials and Matrix Polynomials*, PhD thesis, The University of Western Ontario, London, ON, Canada, 2006.
[2] A. Amiraslani, R. M. Corless, and P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*, IMA J. Numer. Anal., 29 (2009), pp. 141–157.
[3] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, SIAM, Philadelphia, PA, third ed., 1999.
[4] E. N. Antoniou and S. Vologiannidis, *A new family of companion forms of polynomial matrices*, Electron. J. Linear Algebra, 11 (2004), pp. 78–87.
[5] S. Barnett, *Congenial matrices*, Linear Algebra Appl., 41 (1981), pp. 277–298.
[6] J.-P. Berrut and L. N. Trefethen, *Barycentric Lagrange interpolation*, SIAM Rev., 46 (2004), pp. 501–517.
[7] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), pp. 7:1–7:28.
[8] A. Bultheel and M. Van Barel, *Vector orthogonal polynomials and least squares approximation*, SIAM Journal on Matrix Analysis and Applications, 16 (1995), pp. 863–885.
[9] R. M. Corless, *Generalized companion matrices in the Lagrange basis*, in Proceedings EACA, L. Gonzalez-Vega and T. Recio, eds., June 2004, pp. 317–322.
[10] R. M. Corless and N. Fillion, *A Graduate Introduction to Numerical Methods: From the Viewpoint of Backward Error Analysis*, Springer, 2013.
[11] R. M. Corless, N. Rezvani, and A. Amiraslani, *Pseudospectra of matrix polynomials that are expressed in alternative bases*, Math. Comput. Sci., 1 (2007), pp. 353–374.
[12] F. De Terán and F. M. Dopico, *Sharp lower bounds for the dimension of linearizations of matrix polynomials*, Electron. J. Linear Algebra, 17 (2008), pp. 518–531.

[13] F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations and the recovery of minimal indices*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2181–2204.
[14] ———, *Spectral equivalence of matrix polynomials and the index sum theorem*, Linear Algebra Appl., 459 (2014), pp. 264–333.
[15] C. Effenberger and D. Kressner, *Chebyshev interpolation for nonlinear eigenvalue problems*, BIT, 52 (2012), pp. 933–951.
[16] H.-Y. Fan, W.-W. Lin, and P. Van Dooren, *Normwise scaling of second order polynomial matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 252–256.
[17] R. T. Farouki and V. T. Rajan, *On the numerical condition of polynomials in Bernstein form*, Comput. Aided Geom. Design, 4 (1987), pp. 191–216.
[18] M. Fiedler, *Expressing a polynomial as the characteristics polynomial of a symmetric matrix*, Linear Algebra Appl., 141 (1990), pp. 265–270.
[19] ———, *A note on companion matrices*, Linear Algebra Appl., 372 (2003), pp. 325–331.
[20] S. Fortune, *An iterated eigenvalue algorithm for approximating roots of univariate polynomials*, J. Symbolic Comput., 33 (2002), pp. 627–646.
[21] S. Gaubert and M. Sharify, *Tropical scaling of polynomial matrices*, in Positive systems, R. Bru and S. Romero-Vivó, eds., vol. 389 of Lecture Notes in Control and Information Sciences, Springer, 2009, pp. 291–303.
[22] I. Gohberg, M. Kaashoek, and P. Lancaster, *General theory of regular matrix polynomials and band Toeplitz operators*, Integral Equations Operator Theory, 11 (1988), pp. 776–882.
[23] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*, vol. 52, SIAM, 1982.
[24] I. Good, *The colleague matrix, a Chebyshev analogue of the companion matrix*, Q. J. Math., 12 (1961), pp. 61–68.
[25] L. Grammont, N. J. Higham, and F. Tisseur, *A framework for analyzing nonlinear eigenproblems and parametrized linear systems*, Linear Algebra Appl., 435 (2011), pp. 623–640.
[26] N. J. Higham, *The numerical stability of barycentric Lagrange interpolation*, IMA J. Numer. Anal., 24 (2004), pp. 547–556.
[27] N. J. Higham, R. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1218–1241.
[28] P. Lancaster, *Linearization of regular matrix polynomials*, Electron. J. Linear Algebra, 17 (2008), pp. 21–27.
[29] P. Lancaster and P. Psarrakos, *A note on weak and strong linearizations of regular matrix polynomials*, tech. rep., Numerical Analysis Report 470, Manchester Centre for Computational Mathematics, 2005.
[30] ———, *On the pseudospectra of matrix polynomials*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 115–129.
[31] P. W. Lawrence, *Eigenvalue Methods for Interpolation Bases*, PhD thesis, The University of Western Ontario, 2013.
[32] P. W. Lawrence, *Fast reduction of generalized companion matrix pairs for barycentric Lagrange interpolants*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1277–1300.
[33] D. Lemonnier and P. Van Dooren, *Balancing regular matrix pencils*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 253–263.
[34] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.
[35] V. Mehrmann and D. S. Watkins, *Polynomial eigenvalue problems with Hamiltonian structure*, Electron. Trans. Numer. Anal., 13 (2002), pp. 106–113.
[36] C. B. Moler and G. W. Stewart, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256.
[37] B. Parlett and C. Reinsch, *Balancing a matrix for calculation of eigenvalues and eigenvectors*, Numerische Mathematik, 13 (1969), pp. 293–304.
[38] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.
[39] F. Tisseur and N. J. Higham, *Structured pseudospectra for polynomial eigenvalue problems, with applications*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 187–208.
[40] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.
[41] A. Townsend, V. Noferini, and Y. Nakatsukasa, *Vector spaces of linearizations for matrix polynomials: A bivariate polynomial approach*, MIMS EPrint 2012.118, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Dec. 2012.
[42] R. Van Beeumen, W. Michiels, and K. Meerbergen, *Linearization of Lagrange and Hermite interpolating matrix polynomials*, IMA J. Numer. Anal., (2014).